Amendments to the Specification

Please amend the specification as follows.

Kindly replace the paragraph beginning at line 28 of page 1 with the following:

Fig. 2 presents a block diagram of the principal functions of the transmitting device 102 and the base station 104 in a DTX system. A ~~speaker=s~~ speaker's voice is received by an audio input port (AIP) 122 where the voice signal is digitally sampled at some frequency fs, typically fs = 8 kHz. The sampled signal is usually divided into frames of length 10 msec or so (i.e., 80 samples) prior to further processing. The frames are input to a voice activity detector (VAD) 124 and a speech encoder 126. As is known to those skilled in the art, in some devices, the VAD 124 is integrated into the speech encoder 126, although this is not a requirement in prior art systems. In any event, the VAD 124 determines whether or not speech is present and, if so, sends an active signal to the ~~handset=s~~ handset's control interface 128. The ~~handset=s~~ handset's control interface 128 sends a traffic channel request over the control channel 130 to the traffic channel manager 132 resident in the base station 104. In response to the request, the traffic channel manager 132 eventually sends back a traffic channel grant to the handset=s control interface 128, using the control channel 130. Upon receiving the traffic channel grant, the ~~handset=s~~ handset's control interface notifies the VAD 124, the speech encoder 126 and/or the ~~handset=s~~ handset's bit-stream transmitter 134 that a traffic channel 136 has been allocated for transmitting voice data. When this happens, the speech encoder 126 encodes the speech frames and sends the encoded speech signal to the ~~handset=s~~ handset's bit-stream transmitter 134 for transmission over the traffic channel 136 to the appropriate bit-stream receiver 138 associated with the base station 104. In some devices, the speech encoder 126 prepares frames for transmission and sends these to the bit-stream transmitter, whether or not there is voice information to be transmitted. In such case,

the transmitter does not transmit until it receives a signal indicating that the traffic channel 136 is available.

Kindly replace the paragraph beginning at line 20 of page 2 with the following:

In the above-described conventional system, there is delay between the time that frames emerge from the audio input port and the bit-stream transmitter 134 begins to transmit voice data. The overall delay includes a first delay associated with the time that it takes the VAD to detect that voice activity is present and notify the ~~handset=s~~ handset's control interface prior to the traffic channel request, the [[A]]VAD delay[[@]], and a second delay associated, with the time between the traffic channel request and the traffic channel grant, the [[A]]channel access delay[[@]]. The length of the VAD delay is fixed for a given handset, and depends on such things as the frame length being used. The length of the channel access delay, however, varies from talkspurt to talkspurt and depends on such factors as the system architecture and the system load. For example, in the wireless voice over EDGE (Enhanced Data for GSM Evolution) system, the channel access delay is approximately 60 msec, and possibly more. Conventionally, mitigating any type of access delay entails either a) buffering the voice bit-stream until permission is granted, and thereby retarding transmission by that amount of time, b) throwing away speech at the beginning of each utterance ([[A]]i.e., [[A]]front-end clipping[[@)]] until permission is granted, or c) a combination of the two approaches. The buffering option introduces delay, which is detrimental to the dynamics of interactive conversations. Indeed, adding 120 msec of round trip delay just for access delay can break the overall delay budget for the system. The front-end clipping option often cuts off the initial consonant of each utterance, and thus hurts intelligibility. Finally, combining the two options such that less clipping occurs at the expense of delay is less than satisfactory because such an approach suffers from the disadvantages of both.

Kindly replace the paragraph beginning at line 13 of page 3 with the following:

The present invention is directed to a method and system for removing access delay during the beginning of each utterance as the talkspurt progresses. This is done by time-scale compressing, i.e., speeding up, the speech at the start of a talkspurt before it is passed to the speech coder. The speech is speeded up by buffering each talkspurt, estimating the ~~speaker=s~~ speaker's pitch period, and then deleting an integer number of pitch ~~period=s~~ periods worth of speech from the buffered talkspurt to produce a compressed talkspurt. The compressed talkspurt is then encoded and transmitted until the access delay has been fully mitigated, after which the incoming voice signal is passed through without further compression for the remainder of the talkspurt.

Kindly replace the paragraph beginning at line 26 of page 4 with the following:

The VAD 152 outputs an active signal, which indicates an inactive-to-active transition, both to the ~~handset=s~~ handset'scontrol interface 164 and the ADR 154, thereby signifying that voice frames are present. The handset=s control interface 164, in turn, informs the traffic channel manager 166 via the control channel 168 that a traffic channel is needed to send the bit-stream. The traffic channel manager 166, in turn, locates and allocates an available traffic channel and, after the access delay, Da, informs the handset=s control interface 164 by sending an appropriate message back over the control channel 168, which is sent on to the ADR 154. The traffic channel is requested and assigned by the traffic channel manager 166 at the start of each talkspurt. At the end of each talkspurt, the VAD 152 detects that no further speech is being generated, and sends an appropriate signal to the ~~handset=s~~ handset's control interface 164 which, in turn, informs the traffic channel manager 166 that the assigned traffic channel is no longer needed and now may be reused.

Kindly replace the paragraph beginning at line 3 of page 7 with the following:

After the talkspurt is over, an active-to-inactive transition occurs in the VAD 152 and the VAD 152 sends an inactive signal to the ~~handset=s~~ handset's control interface 164. When the ~~handset=s~~ handset's control interface 164 receives and processes the inactive signal, this ultimately results in the traffic channel 160 being freed for reuse by the base station 142. The ~~handset=s~~ handset's control interface 164 then waits for another active signal from the VAD 152, in response to another talkspurt. However, if the talkspurt is very short, e.g., less than the time period T of 500 msec, the system may not have enough time to completely remove the access delay. In this case, the bit-stream transmitter 158 informs the ~~handset=s~~ handset's control interface 164 that there is still data to send, which may defer freeing the traffic channel 160 until all the encoded packets have been transmitted.

Kindly replace the paragraph beginning at line 25 of page 8 with the following:

In the present example, a general purpose VAD based on signal power, such as that described in U.S. Patent No. 5,991,718, is used. The first few active speech frames from this VAD are placed in buffer associated with the ADR and, for various reasons, are not time-compressed, but rather are sent on to the speech encoder. When the transmission channel is granted, the obtained access delay Da is measured and converted to samples. At a sampling rate of 8 kHz, a simulated access delay Da = 60 msecs corresponds to a total of 480 samples that must be removed over the time-scaling interval T = 500 msec. This calls for a speed-up rate r = 0.12 = 60 msec / 500 msec. Since there are 25 frames of length F = 20 msecs in a 500 msec time interval, on average, 480/25 = 19.2 samples should be removed from each frame. To ensure that the cutting process is [[A]]on track[[@]], two accumulators are kept. One accumulator, called target count Tc, keeps track of how many samples should have been removed by the time the current frame is transmitted. Tc is initially 19.2 (since by the time the first frame is sent, about 19.2 samples should have been cut) and is incremented by 19.2 with each passing frame. The second accumulator, called the remaining count Rc, keeps

track of how many more samples must be removed to get rid of the entire access delay. Therefore, in the present simulation, Rc is initially set to 480, and then decreases, each time samples are cut from a frame during the processing.